

# Automating Trajectory Prediction Performance Analyses for the FAA Traffic Management Advisor

Gabriele Enea,<sup>1</sup> Robert A. Vivona,<sup>2</sup> and Vincent Kuo<sup>3</sup>  
*Engility Corporation, Billerica, MA, 01821*

Karen T. Cate<sup>4</sup> and Michelle M. Eshow<sup>5</sup>  
*NASA Ames Research Center, Moffett Field, CA, 94035*

This paper presents a new approach for analyzing the trajectory prediction performance of the FAA's Traffic Management Advisor (TMA). TMA is a deployed system that generates scheduled time-of-arrival constraints for en-route air traffic controllers in the US. The new automated analysis provides a repeatable evaluation of the current trajectory performance metrics, for new releases of TMA, in different traffic and airspace environments, and for current traffic situations. Using a wider set of data, it provides also a higher level of understanding on the causes of possible degradation of the trajectory prediction performance of TMA. The bulk of the work consisted of the development of the ability to filter flights not impacted by controller intervention. Identifying interrupted flights from recorded data is challenging but necessary for a fair and accurate performance test. Currently, no method for identification of flights exists other than a manual review of voice communications. The automated approach was tested with two data sets, from 2006 and 2013. The 2013 data consisted of 24 hours of traffic arriving into Dallas Fort Worth Airport. The results of the testing on this data set showed that the approach selected a statistically significant number of flights to validate the TMA trajectory predictor's performance against the system requirements. New metrics for the evaluation of the TMA trajectory predictor's performance are introduced and compared with the current set of metrics used by the FAA.

## I. Introduction

TRAJECTORY-BASED Operations (TBO) is a cornerstone concept of modernization efforts for air transportation systems worldwide.<sup>1,2, 3</sup> To implement TBO in the short term,<sup>4,5</sup> accurate estimates of times at the meter fix are necessary.<sup>6,7</sup> This function in the US National Airspace System (NAS) is provided by the Traffic Management Advisor (TMA). TMA is now deployed in all 20 en-route Air Route Traffic Control Centers (ARTCCs) and many major Terminal Radar Approach Control (TRACON) facilities in the US. It provides controllers with the time-based metering function and arrival flow visualization.

TMA relies on an internal trajectory predictor, the Trajectory Synthesizer (TS),<sup>8,9,10</sup> to predict the future trajectory of flights. The accuracy of TMA trajectory predictions governs the efficacy of time-based metering operations. Existing procedures to verify and validate TMA trajectory predictions are not applicable to current traffic data and not easily repeatable. This paper describes a new method used to automate the testing techniques in order to improve the testing of new versions of the software and to create a repeatable test that can be performed in a lab environment before any new releases get deployed in the field. The new test is quicker, more flexible, applicable to current traffic data, and gives a higher level of insight into the possible causes of degradation of the trajectory prediction performance of TMA. The most challenging part of this research work was the automation of a filtering algorithm to select the flight set to test the trajectory prediction performance requirement of TMA.

<sup>1</sup> Principal Research Engineer, 300 Concord Road, Suite 400, AIAA Senior Member.

<sup>2</sup> Chief Research Engineer, 300 Concord Road, Suite 400, AIAA Associate Fellow.

<sup>3</sup> Senior Research Engineer, 300 Concord Road, Suite 400, AIAA Senior Member.

<sup>4</sup> Computer Engineer, Mail Stop 210-8, AIAA Senior Member.

<sup>5</sup> Computer Engineer, Mail Stop 210-8.

This paper describes first the problem that the FAA and NASA are facing with the trajectory prediction performance requirements of TMA. The two current approaches to test these requirements are presented next with a discussion of the limitations of the approaches. The central part of the paper is dedicated to the description of the techniques and algorithms used to automate the current test approaches. Proposed additional metrics for the evaluation of the performance of TMA are briefly introduced as possible future work. The paper concludes with discussions of lessons learned from the current work.

## II. Background

In the continuing effort to improve TMA's performance, the FAA, working with NASA, has implemented a performance test that analyzes the evolution of the accuracy of the ETAs calculated by TMA. Since the accuracy of TMA's schedule is tightly related to the performance of its Trajectory Predictor (TP), a meaningful analysis of the TMA system as a whole has to look at both aspects. The objective of the work described in this paper is to use recorded data to identify problems with the TP performance of TMA.

The FAA uses two approaches to test a new release of the TMA system, one in a laboratory environment before the deployment, and one in the field, after a new release of TMA has been deployed. Both approaches have limitations, mainly in time-efficiency.

One of the basic capabilities of TMA is the computation of the un-delayed ETA to the meter fix and runway for each arrival aircraft in the airspace analyzed. Based on these ETAs, TMA computes the sequences and scheduled times of arrival (STAs) to the meter fix and runway for each aircraft to meet the sequencing and scheduling constraints entered by the user.<sup>11,12</sup> The ETAs are calculated based on the four-dimensional (4D) trajectories predicted for each aircraft. The 4D trajectories consist of three spatial dimensions plus time. The accuracy of the predicted 4D trajectories directly impacts the accuracy of the ETAs. Therefore a thorough analysis of the causes of ETA inaccuracies has to look also at the performance of TMA's TP.

To calculate the future trajectory of an aircraft, the TP needs to know the current position, the wind field and the intent of an aircraft. Upon each RADAR sweep, TMA predicts a new trajectory and a new ETA. Similarly, upon any change in the aircraft intent, e.g. if ATC issues an advisory or the pilot applies some control to maneuver or change the speed of the aircraft, TMA predicts a new trajectory with a new ETA.

The current TMA performance requirement evaluated by the FAA focuses on the accuracy of the predicted ETAs for a sample of flights not impacted (interrupted) by any ATC intervention. Any ATC intervention that is not included in a flight plan amendment or that is exchanged via voice communications between ATC and cockpit, is not known by the automation system. Leaving the flights with these unknown interventions in the test sample would not be appropriate to assess the performance of TMA. For this reason these flights have to be identified and removed. Identifying interrupted flights from recorded data is challenging but necessary for a fair and accurate performance test. An automated approach to identify and remove these flights is therefore necessary to apply the performance test to any set of flights.

An uninterrupted flight can be defined as a flight that has not received any clearance from ATC, either lateral, vertical or speed-related, that will change its objectives once it is beyond the freeze horizon. The freeze horizon is the point after which TMA no longer automatically adjusts the STAs for the flights in the schedule. Lateral interruptions can be caused by a "direct-to" clearance, route changes in flight plan amendments, vectoring, or similar maneuvers issued by ATC. A vertical interruption is caused by any interim altitude level-off during a climb or descent or a change in the aircraft's current cruise altitude. A speed interruption is caused by a climb, cruise or descent speed clearance during any of these phases of flight. Any of these interventions may be applied to maintain separation with other traffic to the meter fix where traffic is handed off from the Center controller to the TRACON controller.

A significant part of the work presented in this paper reports on the effort of automating the filtering approach to identify uninterrupted flights from recorded data. Currently, no method for identification of flights exists other than a manual review of voice communications. As a result of this, the FAA uses a static set of flights, recorded in 2006 and manually filtered, to test TMA's accuracy in predicting the ETAs. The use of this data is presented in the next section.

## III. Current Analysis Approach

This section presents the two approaches currently used by the FAA to test the TP performance requirements of TMA.

### A. Data-driven Approach (“Cone” Test)

The performance requirement test, referred to as the “cone” test, uses a data set recorded in 2006 in which a set of flights have been hand-picked to exclude those that were impacted by ATC clearances. The test is run on this fixed set of flights every time a new release of TMA is completed. The performance tested with this approach is the cumulative accuracy of TMA in predicting the ETAs of a set of flights after the freeze horizon at 19 minutes from crossing the meter fix. The requirement states that the ETA error should be linearly decreasing from the freeze horizon to the meter fix crossing.

The specific FAA requirements description is as follows: TMA SHALL [0580] compute the track-based ETA for all aircraft type trajectories to within Root Mean Square (RMS) accuracy of +/- 60 seconds when the aircraft time to the CTAS meter fix or meter fix arc is 19 minutes. The error rate SHALL [0581] decrease linearly as:

$$\text{error rate} = (60 \text{ sec} / 19 \text{ min}) \times (\text{time to CTAS meter fix or meter fix arc})$$

where the upper bound for time to CTAS meter fix or meter fix arc is 19 minutes. The RMS Error is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (MFX_{ATA} - MFX_{ETA})_i^2}{N}} \quad (1)$$

Where:

$$N = \text{Number of aircraft going to meter fix}$$

$$MFX_{ATA} = \text{Actual Time that aircraft } i \text{ crosses the meter fix}$$

$$MFX_{ETA} = \text{Estimated Time when aircraft } i \text{ will cross the meter fix when 19 minutes out}$$

This formula is applied to each flight’s ETA calculated by TMA after the flight is within 19 minutes (actual time) from its meter fix. The flights included in the performance test were manually selected to remove any impact by ATC interventions or under any metering operations in place.

### B. Operations-based Approach

The second approach used by the FAA to test the performance of a new release of TMA is performed directly by the air traffic controllers in the deployed facilities. Running TMA on real traffic flights, controllers look at flights arriving into a specific terminal area. While listening to the radio communications between ATC and the cockpit, the controllers take note of the meter fix ETA calculated by TMA at the predicted freeze horizon for all the flights arriving at a specific meter fix for a set time interval. If they do not hear any clearance that will change (interrupt) the aircraft procedure inside the freeze horizon, the controllers take note of the actual meter fix crossing time of the specific flight and compare it to the TMA predicted time at the freeze horizon. This gives them a direct measure of the performance. If the ETA calculated by TMA at the freeze horizon is less than 60 seconds different from the actual crossing time for the manual sample of flights, then the new release is considered to be operating acceptably.

### C. Limitations of the Current Approaches

The current approaches for the evaluation of TMA trajectory prediction performance requirements have some limitations. For example, the “cone” test presented in section III.A, as currently implemented, cannot be performed on other recorded data than the 2006 data, which contains arrivals in the Houston ARTCC. The static nature of the data set carries the risk that 2006 data may not represent current operations in terms of aircraft mix or procedures. Moreover different Centers with different procedures can present specific, and sometimes unique problems. Those cannot be captured with only one data set. Another limitation of this approach is that, since the data analyzed is from 2006, it is necessary to run TMA with 2006 adaptation data. This is because approach procedures are modified over time by the FAA. Running TMA with current adaptation could cause errors in the conversion of the routes of the flights in the schedule. This problem will be extensively discussed in Section V. The only solution to this issue would be for the FAA to maintain legacy adaptation data. Since many changes to TMA are in the adaptation data, the impacts of these changes would not be evaluated if the old adaptation were used.

The “cone” test evaluates the performance of TMA only for uninterrupted flights. Tying the performance only to these flights may not reflect some changes in the behavior of the TMA TP for the interrupted flights. In a previous study<sup>13</sup> change reports performed by the first TMA deployment contractor were analyzed. The changes reported

were implemented to directly respond to some inappropriate behavior that TMA was presenting in building the arrival schedule. Some of these behaviors are very hard to tie to the “cone” test performance metric, and the work presented in section VI is geared directly toward these issues. From the same study another problem with TMA trajectory prediction was the number of trajectory failures. Failures are those instances when, given the radar track and the objective of the aircraft, the TS cannot calculate a predicted trajectory. Similarly, the “cone” test does not define an acceptable level of trajectory failures.

One limitation with the approach presented in section III.B is that, in the way it is currently performed, it is not easily replicated for to a large number of flights, nor easily automated. Moreover since it is a post deployment test, it looks at the performance of TMA already in the field. The approach presented in this paper aims at identifying any issues that would be observable by the controllers before TMA leaves the lab. With the new approach, any undesirable behavior with new releases of TMA would be identified before the change is deployed. To address some of these issues an automated approach has been developed; this approach is presented in the following section.

#### IV. New Automated Analysis Approach

To mitigate some of the issues presented in the section III.C., an automated, repeatable approach to analyze the performance of TMA trajectory predictor that would be applicable to any flight data set and to any airspace in the NAS was developed. To achieve this goal, the first and most challenging task was to automate the down-selection (filtering) of a set of flights that was not impacted by any intervention by ATC after the freeze horizon and before crossing their intended meter fix. A set of un-impacted flights is necessary to test the performance of TMA since the performance requirement, by definition, only applies to uninterrupted flights.

The automated filtering algorithms classify a flight as being either interrupted or uninterrupted using only the inter-process message data recorded by the deployed or candidate TMA release. The goal was to use recorded ETA and Track data to perform the analysis, with other data such as flight plan as necessary. The critical messages for the filtering algorithm were:

- AIRCRAFT\_ETA\_AVERAGED – This is the record of a TMA generated ETA value for the aircraft. The speed data for the analysis also come from this message that is recorded by TMA roughly every 12 seconds.
- IC\_TRACK\_DATA – This is the record of a single received track update. For each track position, Lat/Lon, altitude and time are recorded in this message.
- AC\_CROSSING\_TIME – This is the record that denotes the aircraft’s actual metering fix crossing time.
- IC\_FLIGHT\_PLAN\_ADD – This is the record of a new flight plan. Both the flight plan and the parsed routes\* are recorded in this message.
- AMEND\_FLIGHT\_PLAN\_RA – This is the record of an amended flight plan. Both the flight plan and the parsed routes are recorded in this message.

The adaptation files necessary for the algorithm:

- National\_waypoints and ZFW\_waypoints - These files are used to determine the location of the waypoints in the flight plan and parsed routes.

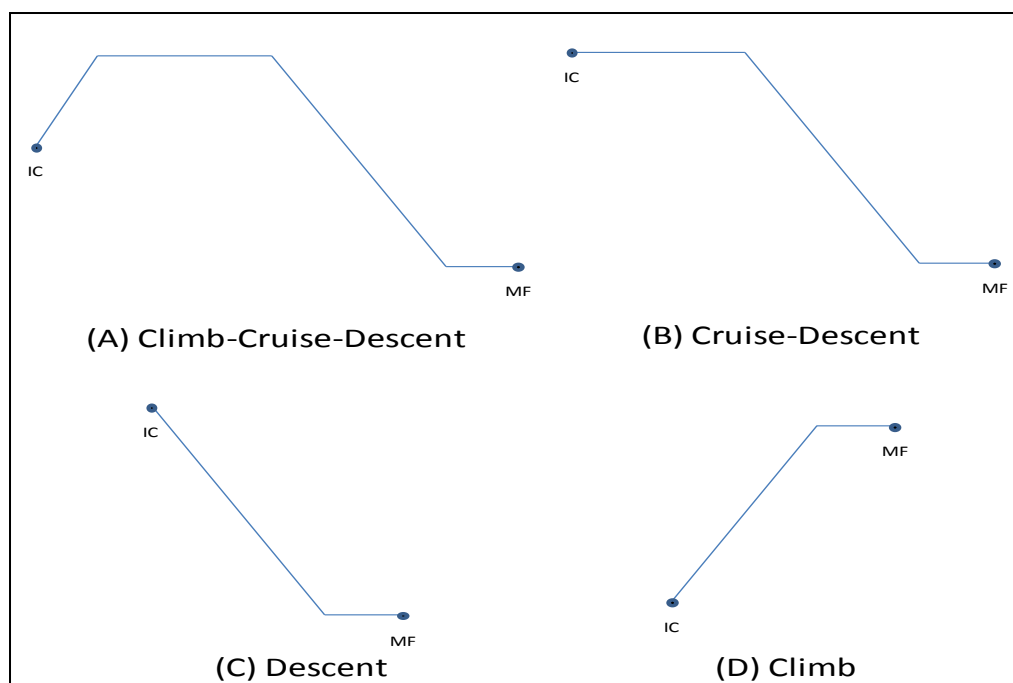
Using these recorded data, the algorithm identifies three types of interruptions: altitude, lateral, and speed. To be defined as interrupted, an aircraft has to change its objective after it has entered the freeze horizon. Therefore, for each aircraft only the data recorded inside the freeze horizon were analyzed by the algorithm.

##### A. Altitude Interruptions

The algorithm to identify altitude interruptions evaluates the vertical profile of the aircraft with respect to pre-defined “acceptable” profiles. There are four possible “altitude profiles” (see Figure 1) representing uninterrupted flights with respect to altitude clearances, although other derivatives of these four are allowed. In Figure 1A, the aircraft’s initial condition at the freeze horizon is in a climb to its cruise altitude. In this case, the aircraft should be uninterrupted in its climb to the cruise altitude and then should stay at this altitude until starting an uninterrupted descent to the metering fix. In Figure 1B, the aircraft’s initial condition at the freeze horizon is established at its cruise altitude and should continue at this altitude until it starts an uninterrupted descent to the metering fix. In Figure 1C, the aircraft’s initial condition at the freeze horizon is already in a descent. The aircraft should continue this descent uninterrupted to the metering fix. In Figure 1D, the aircraft’s initial condition at the freeze horizon is in a climb, but the aircraft will only climb to the metering fix crossing altitude. This climb should be uninterrupted. In all four cases, it is possible that there will be a level-off at the metering fix altitude prior to crossing the metering fix location, as shown in the figures.

---

\* The parsed routes are commonly called AK routes.



**Figure 1 Uninterrupted altitude profiles.**

By examining the four figures, it can be seen that, given the state of the aircraft (climbing, level, descending) at the freeze horizon, there are a limited number of allowable state changes that can occur prior to the meter fix if the flight is uninterrupted. For example, if the aircraft is initially climbing, to be uninterrupted the aircraft can only (1) arrive at the metering fix (CLIMB), (2) level-off and arrive at the metering fix (CLIMB), (3) level-off and then descend to the metering fix (CLIMB-CRUISE-DESCENT), or (4) descend to the metering fix (CLIMB-CRUISE-DESCENT). The final case is equivalent to a very short level-off segment between the climb and descent. Similarly, if the aircraft is level or descending at the freeze horizon, to be uninterrupted the aircraft can only descend to the metering fix.

To compare the vertical profile of the aircraft to the “acceptable” profiles, the algorithm defines whether a contiguous portion of aircraft tracks (a segment) is climbing, level or descending. The algorithm looks at the maximum altitude difference across a set of tracks for a period of at minimum 60 seconds length. Consistent with this minimum value, climbing, level or descending segments are created to cover the length of the tracks inside the freeze horizon. For the  $N$  tracks in a segment:

- Max altitude,  $h_{\max}$  = largest altitude value
- Time at max altitude,  $t_{\max}$  = time of the track with the maximum altitude
- Min altitude,  $h_{\min}$  = smallest altitude value
- Time at min altitude,  $t_{\min}$  = time of the track with the minimum altitude

The segment (i.e., all tracks within it) is considered LEVEL if

- $h_{\max} - h_{\min} \leq 350$  ft

The segment is considered CLIMBING if

- $h_{\max} - h_{\min} > 350$  ft
- $t_{\max} > t_{\min}$

The segment is considered DESCENDING if

- $h_{\max} - h_{\min} > 350$  ft
- $t_{\max} < t_{\min}$

The choice of 350 feet as a threshold value was purely empirical. It was based on observation of the data, and it was calibrated testing different values and observing the results of the filter. Once the algorithm has identified all of the altitude segments, the changes in state (level, climbing, descending) between the segments are compared to the allowable changes in state based on the profiles in Figure 1. Unallowable state changes are an indication that the aircraft was interrupted by an altitude clearance.

## B. Lateral Interruptions

Similar to the altitude interrupt algorithm, the approach for identifying unallowable lateral state changes when analyzing track data inside the freeze horizon was defined. Lateral state changes are used as an indication of: (1) whether the flight is following the Flight Plan route used by TMA, and (2) whether the aircraft track is in a turn. The relationship to the flight plan route is identified by comparing the track data with the Flight Plan route used by TMA to predict the ETAs at the meter fix. The Flight Plan route is received as a series of waypoints. Depending on the relationship between the Flight Plan route and the track data, the algorithm defines whether the aircraft was interrupted by a lateral clearance such as a vector. When the relationship cannot be unambiguously identified, the flight is dismissed from the ETA accuracy analysis. The approach to identifying a lateral interrupt is summarized in the following sections.

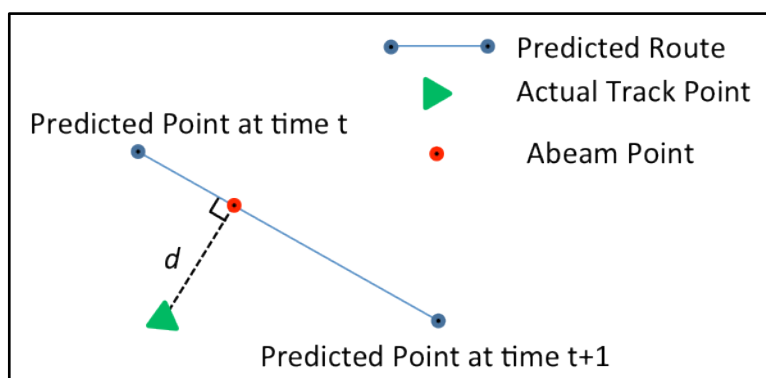


Figure 2 Definition of abeam point

### 1. Preliminary Calculations

To identify if a flight is laterally interrupted, a few preliminary definitions and calculations are necessary:

- Abeam point is defined as the projection, in an XY earth plane, of a track data point, onto the Flight Plan route, as represented in Figure 2.

- For each track data point the abeam point distance  $d$  between the flight plan route and the track is calculated. If the distance  $d$  is greater than 5 nautical miles, the track point is considered out of conformance with

the route and therefore not following its flight plan route. The choice of 5 nautical miles for the  $d$  threshold value here was empirical and based on the observation of the data. The objective of the filter was not to strictly identify flights on their route, but to interpret the objective of the aircraft. This value may be tuned when applied to different datasets.

- To identify Turns, the x/y track values of a series of track data are analyzed to fit a linear polynomial line to the data. The abeam distance between each track data and the line is calculated. If the distance between the line and a track point is greater than 1 nautical mile, the segment is considered to be a turn segment, otherwise it is a straight segment. Similar to the altitude interrupt algorithm, this algorithm uses this definition of straight and turn segments to divide the lateral tracks inside the freeze horizon into a series of straight or turn segments.

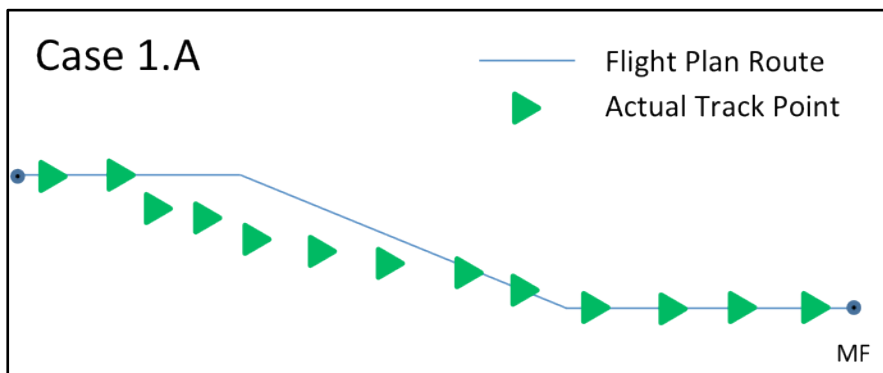
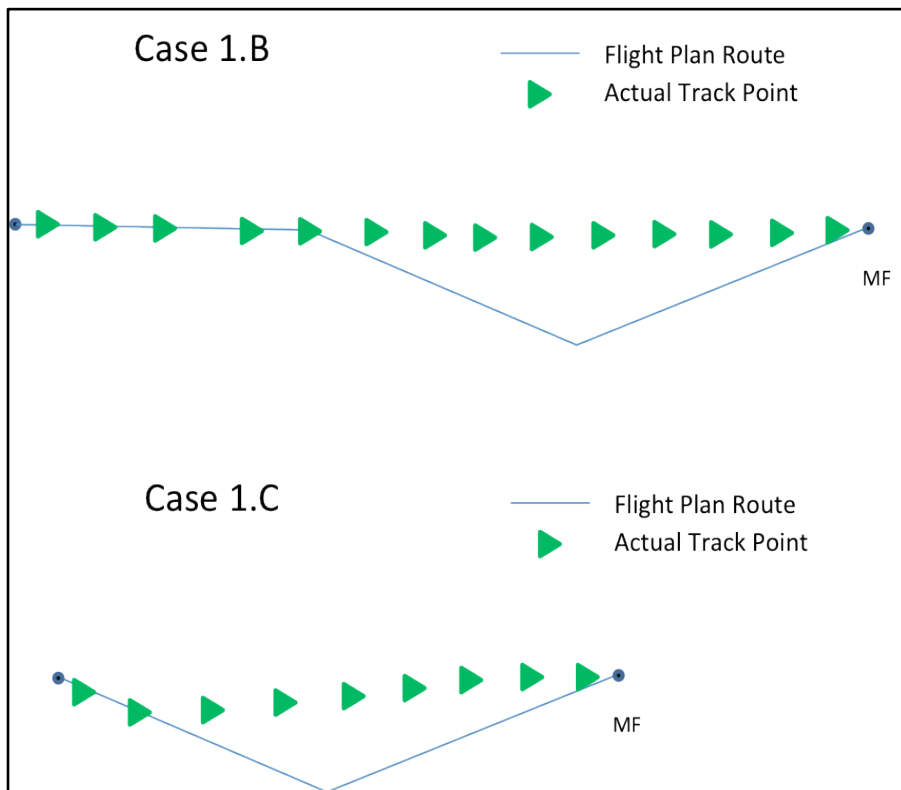


Figure 3 Flight initially following the Flight Plan route

### 2. Lateral Interruptions Algorithm

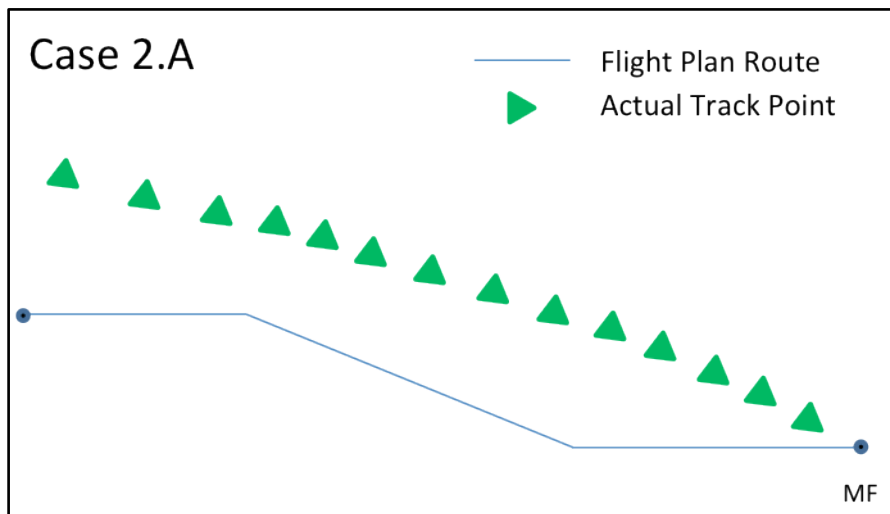
The lateral interrupt algorithm is based on a series of cases representing different relationships between the aircraft tracks and the flight plan route. Each case is described next.



**Figure 4 Special cases: Flight initially following the Flight Plan route. Excluded from analysis.**

horizon, then there are two possibilities:

- If there are no turns after crossing the freeze horizon (Case 1.B), the behavior is ambiguous (interrupt would depend on when the clearance to leave the route was received prior to or after the freeze horizon) so it is excluded from the analysis;



**Figure 5 Flight initially off the Flight Plan route**

If the aircraft is initially following the flight plan route (Case 1.A, Figure 3) the following situations can occur:

- If the tracks follow the route all the way to the meter fix and no route changes that change the lateral path received after the freeze horizon, then the flight is not interrupted;
- If a new flight plan amendment and new route<sup>†</sup> are received after the freeze horizon, then the flight is considered interrupted;
- If the flight leaves the flight plan route (abeam point  $> d$ ) but no flight plan amendment and no new route are received after the freeze

- If there are turns anywhere after the freeze horizon (Case 1.C), then the flight is considered interrupted.

For Case 1.B (Figure 4 top) the ambiguity comes from the fact that it's impossible to determine only from the comparison of the flight plan route and the actual track points if the flight was proceeding direct to the meter fix due to a clearance received before the freeze horizon or if it was originally on its flight plan route and

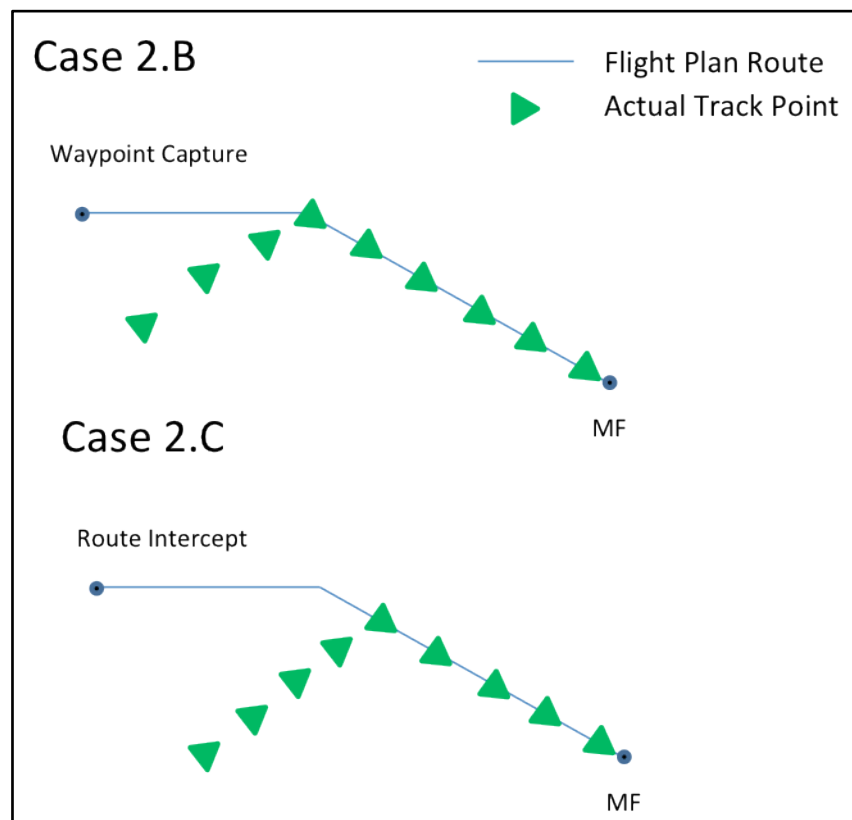
<sup>†</sup> To be considered a "new" route it has to be laterally different from the Flight Plan route.

was cleared direct to the meter fix after the freeze horizon. If the clearance was prior to the freeze horizon, the flight is uninterrupted. If after the freeze horizon, then it was interrupted. Since the difference cannot be determined only from the flight plan route and the actual track points, these flights are flagged and excluded from the analysis. For Case 1.C (Figure 4 bottom), since there is a turn after the freeze horizon that takes the flight off the flight plan route, it is safe to assume that a lateral clearance was issued by the controller after the freeze horizon and therefore the flight is considered interrupted.

If the aircraft is initially off the flight plan route (Case 2.A in Figure 5), then the following cases can occur:

- If no turns after the freeze horizon, the flight is not interrupted
- If any turns after the freeze horizon and the aircraft is NOT on the route when the turn occurs, then the flight is considered interrupted
- If there are turns onto route or to connect to route then flag for ambiguous behavior and exclude from the analysis.

In Case 2.B and Case 2.C (Figure 6), the clearance to connect to the route could have either occurred before the freeze horizon (uninterrupted) or after (interrupted). Since these two cases cannot be separated from just the flight plan and track data, they are excluded from the analysis.



**Figure 6 Special cases: Flight initially off the Flight Plan route. Excluded from analysis.**

is currently flying. “Typical” variations in climb, cruise and descent speed are very different. For this reason the speed observations were segregated by phase of flight. For flights that included a climb phase of flight, speed interruptions during the climb phase were ignored. The speed interrupt algorithm was therefore based on the observation of:

- Speed variations during the cruise phase in terms of Calibrated Air Speed (CAS), and
- Speed variations during the descent phase to identify “typical” CAS evolutions.

These two pieces of information showed clear trends in speed profiles consistent with pilot procedures that are known to be followed today in the data analyzed.

### C. Speed Interruptions

Similar to the previous algorithms, the speed filter has the objective to identify speed controls applied by ATC that would invalidate the flight for use in the analysis. Since these impacts have to be identified from recorded speed data that are calculated by TMA from the positions obtained from RADAR tracks, this part of the analysis was the most challenging. To partially overcome this problem, the speed data were pre-processed to remove obvious outlier data points identified as creating unrealistic speed jumps.

The variations in speed that can be observed inside TMA’s freeze horizon are affected by the phase of flight the aircraft



The speed filter identifies the evolution of the speed inside the freeze horizon. The CAS values are segregated in segments of increasing (acceleration), decreasing (deceleration) and maintain (constant) CAS values. A blend of the approaches used for the lateral and altitude algorithms was used to identify the speed segments.

### 1. Speed Variations During Cruise

Assuming that during cruise at a constant altitude the speed data would be “cleaner”, the speed filter during cruise was only looking for maximum CAS variation. The maximum CAS variation allowed along the cruise segments was 20 knots. If, after the removal of outliers, the variation between the minimum and maximum values is greater than allowed, the flight was considered interrupted.

To protect against common speed changes that occur in the cruise phase of flight when transitioning from the aircraft’s climb speed to cruise speed, a check for this possible non-interrupt top-of-climb (TOC) effect was performed. The algorithm checks to see if the TOC is within or just outside the freeze horizon. If so, an initial speed change starting near the TOC is ignored prior to determining the maximum and minimum cruise CAS values. Similarly, changes to the descent speed profile can occur at the very end of the cruise phase of flight. To remove these non-interrupt impacts, any final acceleration/deceleration segments prior to top-of-descent (TOD) are also ignored when determining the cruise CAS variation.

### 2. CAS Profiles During Descent

A preliminary analysis of descent speed profile plots, like the one presented in Figure 7, showed that multiple target CAS values during descent could be identified. Therefore the descent speed filter was designed to identify if more than one target CAS, i.e. more than one constant speed segment, was identified during the descent phase. The algorithm, like the cruise speed algorithm, decomposes the speed during descent into segments. If more than one maintain speed segment is present, the algorithm considers the flight as interrupted. The difference between the maximum/minimum CAS observed in the first and second maintain segments has to be more than 20 knots to be considered an interrupt. This condition was added to avoid filtering flights that have two (or more) constant segments separated by acceleration or deceleration segments but that are part of the same target speed. This effect is caused by the variability of the descent CAS data.

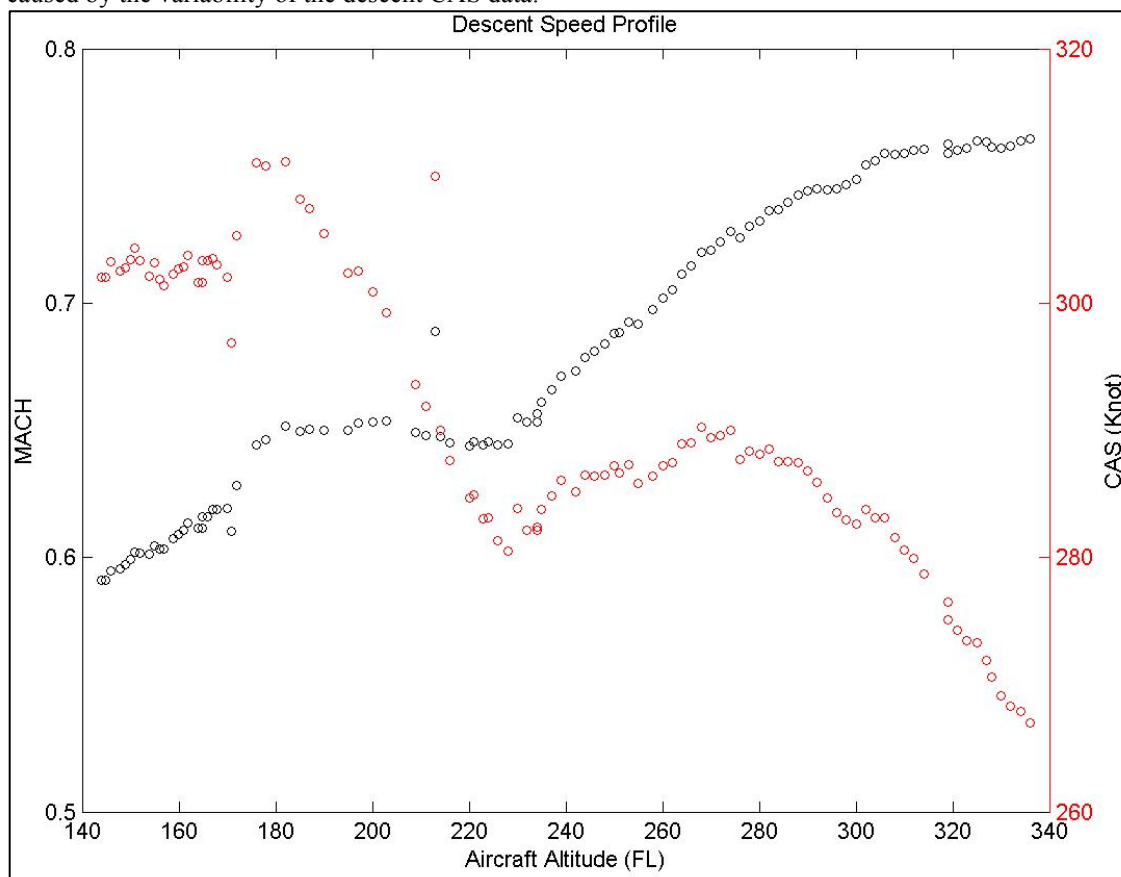


Figure 7 Example of descent speed profile plot

The speed filter was intentionally calibrated not to impact too many flights and to filter only flights with behaviors in which clearly (at least) two CAS values were targeted inside the freeze horizon. This is because it was acknowledged that speed changes, either from ATC issued control or from pilot decisions, are hard to identify from speed data that often are noisy or unstable, especially while aircraft are turning.

## V. Automated Filtering Results

### A. 2006 Data Recording

To validate the automated filtering algorithm, the 2006 data set used by the FAA to run their performance test (see section III.A) was analyzed. The data set included 194 flights manually selected by the FAA. These flights were considered uninterrupted by ATC intervention. The application of the automated filter to this data set disclosed some problems with the use of old data. The current TMA release was run with 2006 track data but with 2010 adaptation data, and this caused some problems in the translation of the flight plans into waypoints. Many flights in the data set were assigned incorrect routes, missing the meter fix they were supposed to cross. An assessment of the performance of TMA using the automated filtering algorithm was therefore meaningless for these flights. A summary of the results of using the automated filter approach on the 2006 data set is presented in Figure 8.

What was interesting was to compare the results of the manual and automated filter on the remaining 93 flights. It must be remembered that all the flights in this data set were considered uninterrupted by the FAA. The automated filter instead identified 40 of these flights as being interrupted for either intermediate altitude level-offs (13 flights), leaving their routes (20 flights), or receiving speed clearances (17 flights). The correctness of these interrupts were verified using plots of the altitude/track/flight plan data. Although the automatic filter could be further tuned, this

means that 40 flights were erroneously included by the manual selections of the uninterrupted flights, one of the limitations of the method already discussed in section III.C.

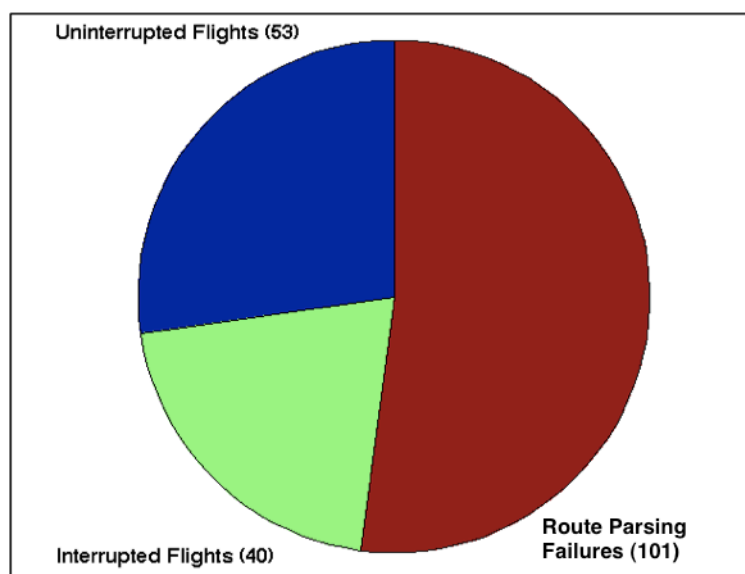


Figure 8 Summary of 2006 data set flights.

### B. 2013 Data Recording

To test its performance with current recorded data, the automated filter was applied to a 24-hour data set from Fort Worth Center (ZFW) recorded on May 8, 2013. The data set included a total of 1,468 meter fix-crossing flights. Of these, 206 were not processed because of either missing routes or missing track or speed data. 759 were identified as interrupted by the automated filter and 503 were uninterrupted (Figure 9-10) and

therefore used for the performance “cone” test (the accuracy performance requirement ([0580] introduced in section III.A).

The results of the automated filtering support the decision to calibrate the speed filter to have the smaller impact on the overall results. The majority of flights are filtered because they leave the route that they are supposed to fly (486 of them, summing all the categories with route interruptions). Discrepancies between the route predicted by TMA and the route actually flown are a known significant cause of inaccuracy in the prediction of the ETAs by TMA. The second most number of flights (328, summing all the categories with altitude interruptions) are removed because inside the freeze horizon they level-off at intermediate altitudes not included in their flight plans. As expected by design, the speed filter impacted by far the smallest number of flights (234, summing all the categories with speed interruptions).

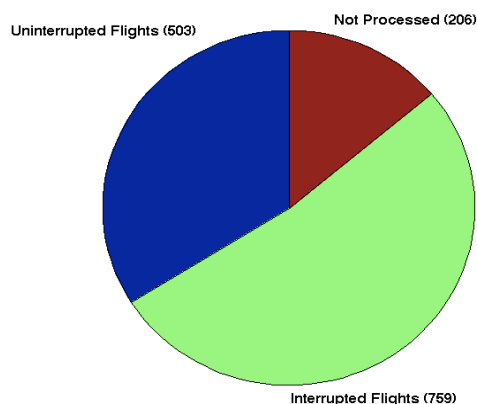


Figure 9 May 2013 Data Set Summary

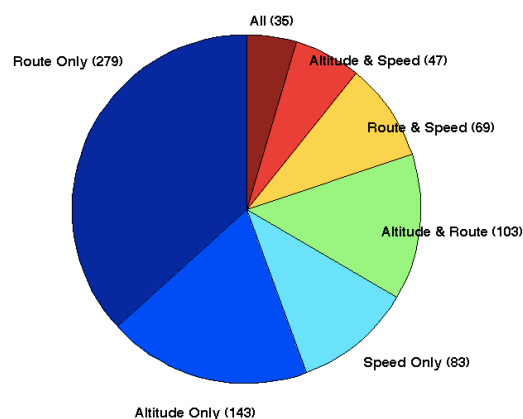


Figure 10 May 2013 Filter Result Summary

Removing these flights from the performance test is necessary from the definition of the performance requirement presented in section III.A. The filtering algorithm presented in this paper does it in an automated fashion. The remaining flights in the data set are then used to test the performance requirement of TMA, and the results for this data set are presented in Figure 11. Although the ETA evolution for some flights are not inside the cone, the majority of the 503 uninterrupted flights have an ETA error at the freeze horizon of less than 60 seconds. As a result the aggregated metric of the RMS error for this data set is 41.48 seconds, well below the performance target of 60 seconds. TMA would pass its performance requirement [0580] for the RMS error.

#### May 13 2013 Uninterrupted Flights RMSE@19 = 41.48 seconds

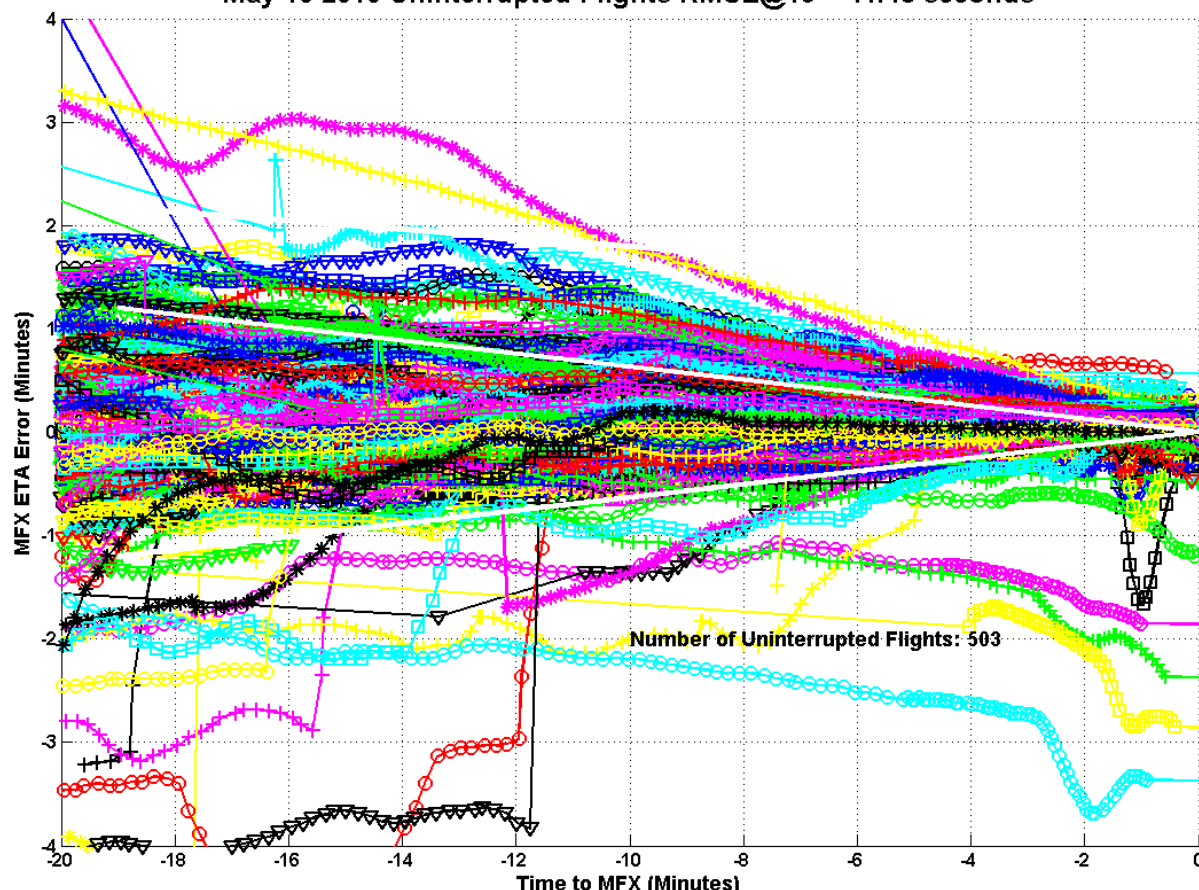


Figure 11 May 2013 data set "cone" test results.

## VI. New TMA Performance Metrics

A previous study<sup>13</sup> on the changes performed to TMA's TP after its deployment in the field showed issues with the predictions of flights impacted by ATC interventions. These flights are not included in the "cone" test presented in section III, representing one of the test's biggest limitations. This section describes an automated analysis tool that calculates proposed new performance metrics for TMA that are not dependent on identifying uninterrupted flights, therefore addressing this limitation.

The objective of the work was to create an automated analysis tool that, calculating new metrics, would support the analysis of the performance of TMA in handling all flights, both impacted and un-impacted by ATC interventions. The tool runs a new approach that focuses on analyzing every ETA generated by TMA and identifies problems regardless of interruptions. A new set of metrics were created based directly on ETA behaviors:

- Failures<sup>†</sup> metric: Total number of failures;
- Successful ETA metrics:
  - ETA jumps<sup>§</sup>, defined as the difference between two consecutive ETAs (Figure 12)
  - ETA creep<sup>§</sup>, defined as the movement of the ETA values over a defined period of time or across a defined number of consecutive prediction cycles (Figure 13).

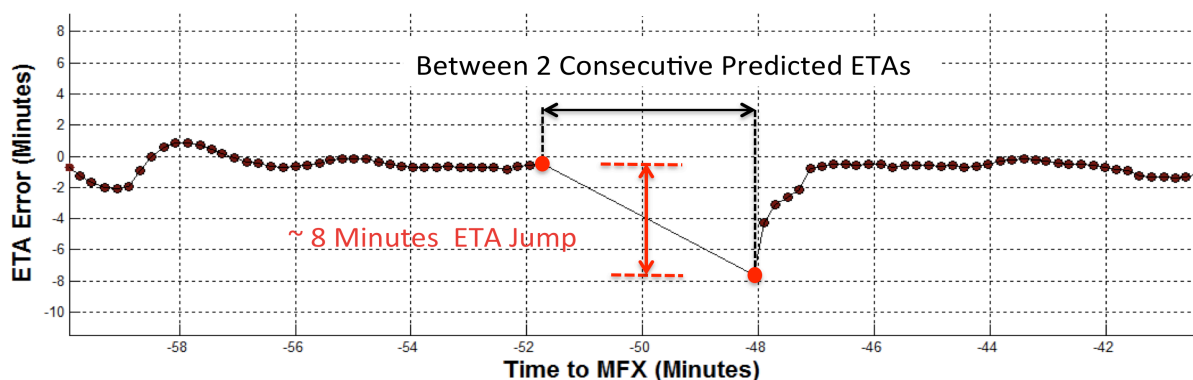


Figure 12 ETA jump metric example

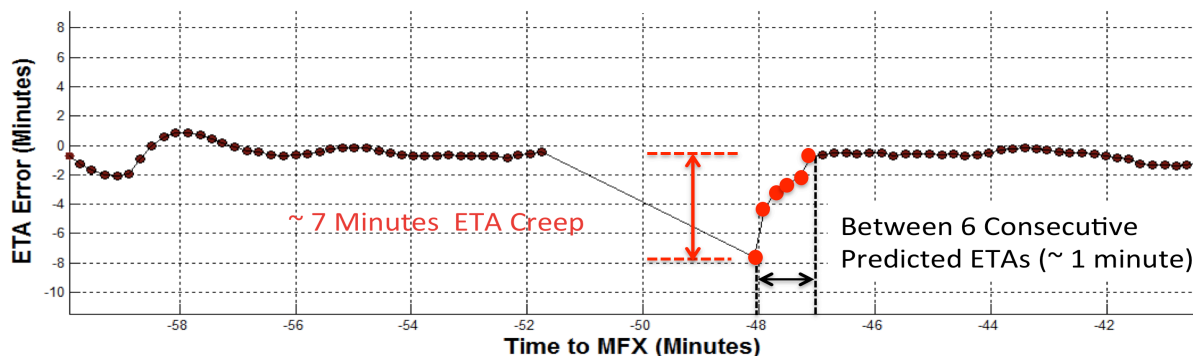


Figure 13 ETA creep metric example

<sup>††</sup> Currently the trajectory failures are not recorded in the data sets analyzed

<sup>§</sup> The terminology is based on terminology directly derived from the change reports collected by the FAA

One of the benefits of these metrics is that they are defined in the “language” of the air traffic controllers, the final users of TMA, which makes the definition of acceptable values easier. Although more effort is necessary to develop and refine them, these metrics represent an enhanced ability to effectively assess TS performance in support of TMA.

The analysis tool starts by looking at all the ETAs created by TMA in the data sample. The ETAs are plotted in a histogram and color-coded according to the prediction type from which they were created as presented in Figure 14. There are four types of prediction in TMA:

- PROPOSED, only for internal departures, the flight has not taken off yet so the ETA is based on the proposed departure time
- DEPARTED, only for internal departures, the flight has taken off but TMA has not received track data for the flight yet
- ESTIMATED, only for external flights, TMA has not received track data for the flight so the ETA is based on the time the flight is estimated to enter the center
- ACTIVE, for all flights, TMA has received track data for the flight and the predicted ETA is based on that information.

The analysis tool allows filtering of the ETA jumps relative to each type of prediction and plotting the desired results for all the flights in a sample. A series of statistics are performed on each category to identify which category presents the biggest ETA jumps. This capability includes all the ETAs calculated by TMA, but also allows down-selecting the flight category of interest for a specific targeted analysis.

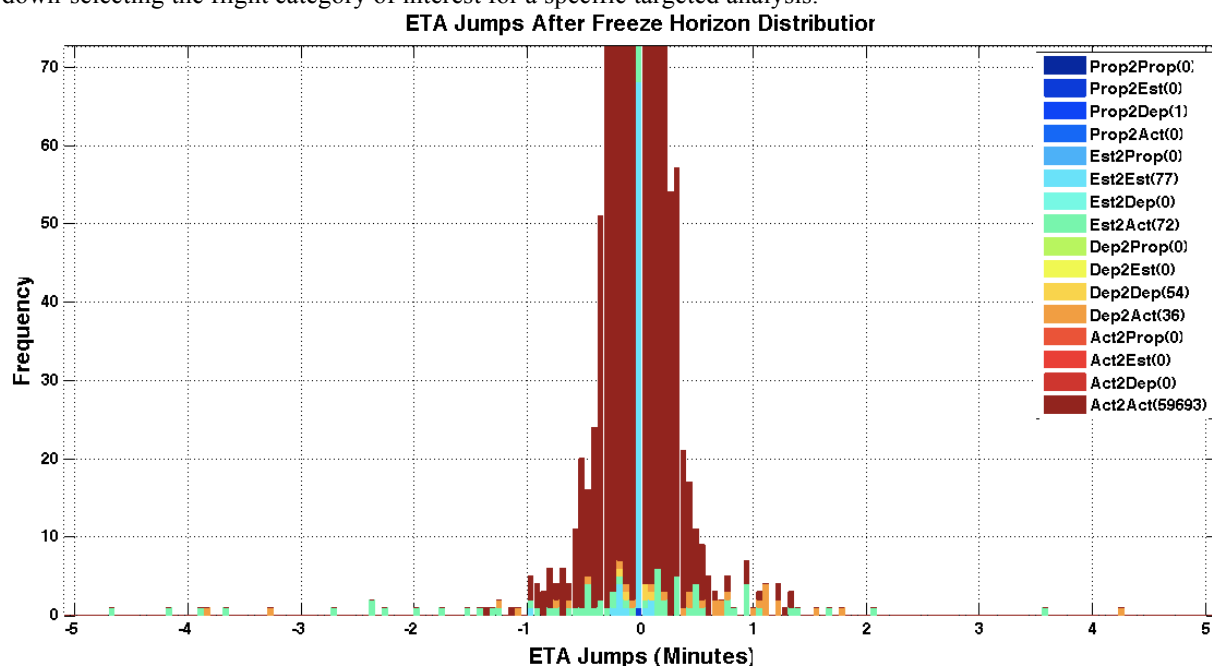


Figure 14 ETA jumps distribution, all prediction types.

Once a specific prediction type and time frame from the meter fix have been chosen, an “acceptable” jump level can be selected to verify if any flight in the data set violates the threshold. In Figure 15 one possible analysis is presented for example. The “acceptable” jump level was set to one minute. A jump of one minute between two ETAs after the freeze horizon is likely to be noticed by the controllers, as the flights’ STAs have been frozen and they are trying to zero the error between ETA and STA. The tool will not only present the number of flights that are violating the pre-determined threshold, but also return the flight IDs to facilitate further analysis. The threshold can be set to any value and adjusted if a different “unacceptable” value is identified.

It is important to mention that some ETA jumps are caused by a change in the intent of the flight that TMA is predicting. Some of these intent changes are recorded in the flight plan as amendments. These amendments are not only easy to identify in the data set, but represent also a legitimate behavior of TMA. Therefore the tool identifies which of these ETA jumps are “legitimate” and removes them from the recorded set of “erroneous” ETA jumps.

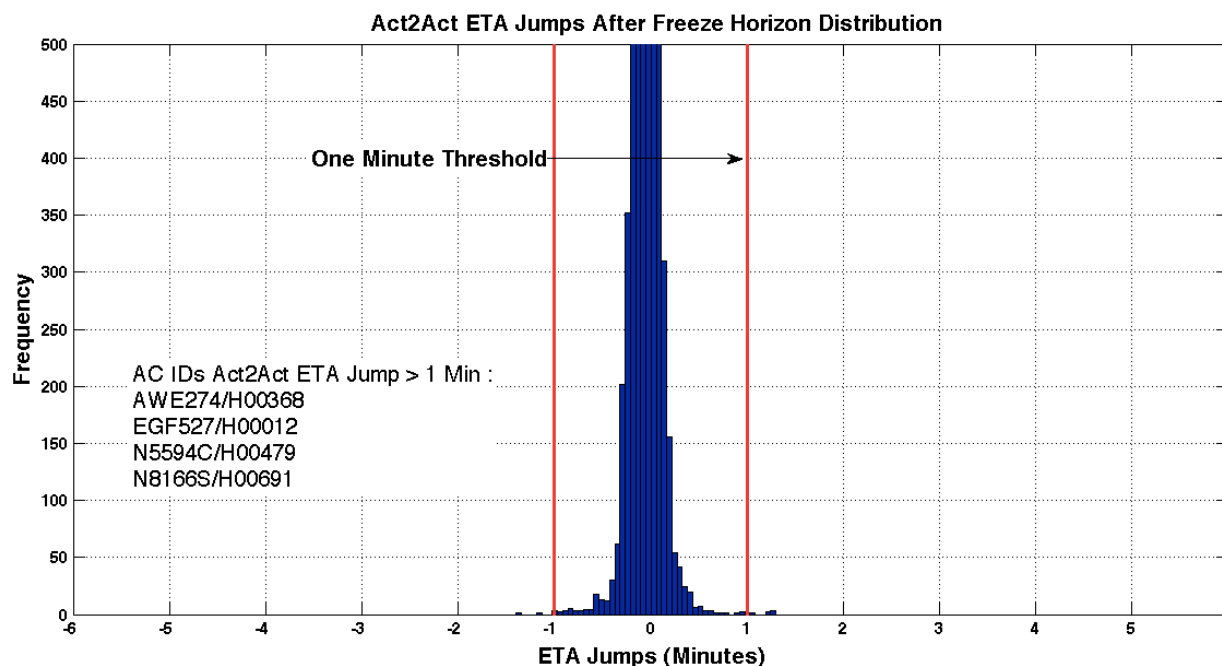


Figure 15 Example of the analysis that can be performed: ETA jumps, ACTIVE predictions only after the freeze horizon.

The tool also allows to closely examine the details of any flight that is violating the ETA jump threshold, plotting a single flight's ETA evolution history. In Figure 16 the detailed plot for a flight, presented by the tool in Figure 15, is shown. On the top part of the plot the ETA error history is presented versus the time to cross the meter fix. The red line on the left of the top plot represents the freeze horizon. In the example, a jump of 62 seconds between two ETAs is highlighted in the plot. On the bottom plot of Figure 16, the histogram of the ETA jumps is presented to complete the data on the flight.

A similar analysis to the one performed on ETA jumps can be performed on ETA creeps. Setting an acceptable level of creep the tool would present the same type of plots that are presented for the jump metric.

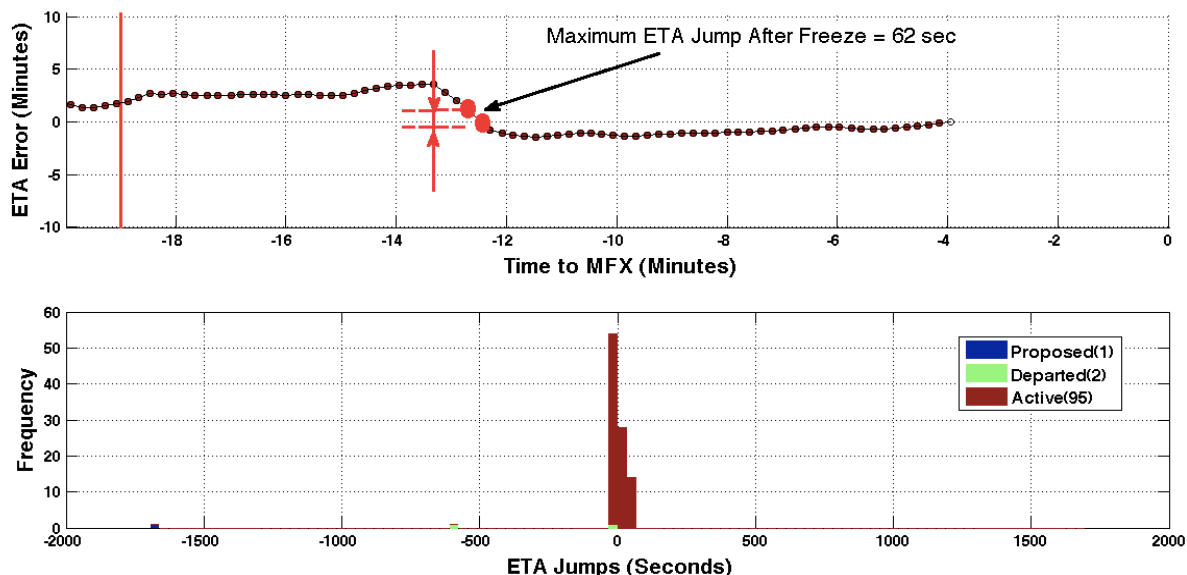


Figure 16 Example of the analysis that can be performed: detailed plot of ETA error and ETA jumps for a single flight.

#### A. Benefits of the New Metrics

The first benefits of the new ETA analysis approach, presented in this section, is that the test analysis is easy to run and provides system-level performance metrics. It also provides the information on the flights that fail the test to



further investigate the causes. The approach can be considered complimentary to the RMS Error metric at 19 minutes test presented in section III and used to corroborate it. If the ETA jump and creep metrics are correctly calibrated they can be directly correlated to the “cone” test. It is also able to assess the performance of TMA with a set of flights including those that are interrupted and maneuvered by the ATC. This is useful since TMA is deployed in the field and has to be able to handle all flights. Lastly, since the metrics are derived directly from controller reported issues, there is a high probability that if TMA passes the ETA jumps and creep tests, its behavior will be acceptable to the controllers.

## VII. Conclusions and Discussion

This paper presented an automated approach to test the trajectory prediction performance of TMA. The approach has the advantage that it does not require a specific set of flight data, but can rather work with data from any site or time. This feature has the potential to accelerate the transition of new releases of TMA from the lab to the field. The application of the performance test to current data assures that the performance of TMA is tested against current traffic with up-to-date operational characteristics. It removes the need to maintain legacy adaptation data necessary to correctly run TMA with old data sets.

Being a laboratory test, it has the advantage, compared to the test described in section III.B, to test the TP performance of TMA before it is actually deployed in the field. This is because the data can be recorded on a test string or in the lab using live traffic feeds. The automated test also helps to discover potential causes of issues by identifying single flights for which the analysis tool creates and stores detailed information.

As an example of the potential of the test, the data presented in section V.B, demonstrate that the approach can be applied to a large set of data, 24-hour recording in this case, and results in a significant sample of uninterrupted flights to apply the performance test required by the FAA. The tool has been tested mainly with two sets of data, one from IAH (the 2006 set) and the large one from DFW (the 2013 set). More data from different centers would be required to make sure that the test can be run without issue for a wider range of operational environments. This could be a possible area for future work.

The set of new metrics presented in section VI have the additional potential to test the performance of TMA with flights that are impacted by controller interventions. Those metrics make no distinction between interrupted and uninterrupted flights. This should be the ultimate goal of the new set of metrics implemented in an automated fashion. Nonetheless more work is necessary to refine these metrics and to make sure that they correctly represent the undesirable behavior of TMA.

## Acknowledgments

The authors would like to thank Alan Lee, Harry Swenson, Todd Farley and Banavar Sridhar at NASA Ames Research Center for their precious comments that significantly improve the quality of this paper.

## References

- <sup>1</sup> FAA, NextGen Implementation Plan 2012, Washington, DC, March 2012. URL: [http://www.faa.gov/nextgen/implementation/media/NextGen\\_Implementation\\_Plan\\_2012.pdf](http://www.faa.gov/nextgen/implementation/media/NextGen_Implementation_Plan_2012.pdf)
- <sup>2</sup> SESAR Joint Undertaking, The European ATM Master Plan, Edition 2, October 2012. URL: <https://www.atmmasterplan.eu/download/25>
- <sup>3</sup> Australian Strategic Air Traffic Management Group (ASTRA). (2007). The ATM Strategic Plan (Part A).
- <sup>4</sup> Mutueal, L. H., Neri, P., Paricaud, E., Initial 4D Trajectory Management Concept Evaluation, 10<sup>th</sup> USA/Europe Air Traffic Management R&D Seminar, Chicago, IL, June 2013.
- <sup>5</sup> Gaydos, T. L., Liao, B., Smith, E.C., Wang, L., Analysis of Conflict Detection Performance for Trajectory-Based Descent Operations, 10<sup>th</sup> USA/Europe Air Traffic Management R&D Seminar, Chicago, IL, June 2013.
- <sup>6</sup> Joint Planning and Development Office, NextGen Concept of Operations, Version 3.1, 2011.
- <sup>7</sup> Haraldsdottir, A., Scharl, J., Berge, M., Shoemig, E., Coats, M., Arrival Management with Required Navigation Performance and 3D Paths, 7<sup>th</sup> USA/Europe Air Traffic Management R&D Seminar, Barcelona, Spain, July 2007.
- <sup>8</sup> Erzberger, H., Tobias, L., A Time-Based Concept for Terminal Area Traffic Management. *Proceedings of the 1986 AGARD Conference, No. 410 on Efficient Conduct of Individual Flights and Air Traffic*, Brussels, Belgium, 1986.
- <sup>9</sup> Slaterry, R., Zhao, Y., En-Route Descent Trajectory Synthesis for Air Traffic Control Automation. *American Control Conference*, Seattle, WA, June 21-23, 1995.
- <sup>10</sup> Slaterry, R., Zhao, Y., Trajectory Synthesis for Air Traffic Automation, *Journal of Guidance, Control and Dynamics*, Vol. 20, No. 2, Mar.-Apr. 1997, pp 232-238.

<sup>11</sup> NASA Aviation System Division, *Traffic Management Advisor Overview*, URL: <http://www.aviationsystemsdivision.arc.nasa.gov/research/foundations/tma.shtml>, accessed January 2013.

<sup>12</sup> Swenson, H. N., Hoang, T., Engelland, S., Vincent, D., Sanders, T., Sanford, B., and Heere, K., Design and Operational Evaluation of the Traffic Management Advisor at the Fort Worth Air Route Traffic Control Center, *1<sup>st</sup> USA/Europe Air Traffic Management R&D Seminar*, Saclay, France, June 1997.

<sup>13</sup> Enea, G., Vivona, R.A., CTAS Trajectory Interoperability Experiment – Final Report, Engility Corporation, Billerica, MA, September, 2012.